MANIPULATION OF ORGANIC ACID BIOSYNTHESIS AND SECRETION

The present invention relates to nucleic acid fragments encoding amino acid sequences for organic acid biosynthetic enzyme polypeptides in plants, and the use thereof for the modification of organic acid biosynthesis and secretion in plants. In particularly preferred embodiments, the invention relates to the combinatorial expression of malate dehydrogenase (MDH) and/or phosphoenolpyruvate carboxylase (PEPC) and/or citrate synthase (CS) in plants to modify organic acid biosynthesis and secretion.

Documents cited in this specification are for reference purposes only and their inclusion is not acknowledgment that they form part of the common general knowledge in the relevant art.

Organic acids, such as citrate and malate, are key metabolites in plants. They are involved in numerous processes, including C4 and Crassulacean acid metabolism (CAM) photosynthesis, stomatal and pulvinual movement, nutrient uptake, respiration, nitrogen assimilation, fatty acid oxidation, and providing energy to bacteroids in root nodules. For example, malate plays a key role in root nodule metabolism and nitrogen fixation, serving as the primary carbon source for bacteroid maintenance and nitrogenase activity, and is also tightly linked to nodule nitrogen assimilation. Furthermore, the complexing role of organic acids produced and excreted from plant roots has also been associated with tolerance to the aluminium cation Al^{3+} which is toxic to many plants at micromolar concentrations. Aluminium toxicity has been recognized as a major limiting factor of plant productivity on acidic soils, which account for approximately 40% of the earth's arable land.

The tricarboxylic acid cycle (TCA), also known as Krebs cycle (after its discoverer Hans Krebs) or citric acid cycle, moves electrons from organic acids to the oxidized redox cofactors NAD^+ and FAD, forming NADH, $FADH_2$, and carbon dioxide (CO_2). The reaction sequence of the TCA cycle involves: in a reaction catalysed by citrate synthase (CS), acetyl-CoA formed by the pyruvate dehydrogenase complex combines with oxaloacetate to produce the C_6 tricarboxylic acid, citrate. In the overall cycle, the citrate is oxidized to produce two molecules of CO_2 in a series of reactions that leads to the formation of one oxaloacetate, three NADH, one $FADH_2$, and one ATP. The resulting oxaloacetate

reacts with another molecule of acetyl-CoA to continue the cycle. The oxidative decarboxylation of pyruvate yields an additional CO_2 and NADH. Thus the TCA cycle brings about the complete oxidation of pyruvate to three CO_2 plus 10 electrons, which are stored temporarily as 4 NADH and 1 $FADH_2$.

Cytosolic reactions generate products that are transported into the mitochondria to feed the TCA cycle. The nature of the end product of the glycolytic reactions in the cytosol of plants is determined by the relative activities of the three enzymes that can utilize phosphoenol-pyruvate (PEP) as substrate. Both pyruvate kinase and PEP-phosphatase form pyruvate; while PEP-carboxylase (PEPC) generates oxaloacetate. Pyruvate is transported directly into the mitochondrion. Oxaloacetate is either transported directly into the mitochondrion or first reduced to malate by cytosolic malate dehydrogenase (MDH).

Before entering the TCA cycle proper, pyruvate is oxidised and decarboxylated by the pyruvate dehydrogenase enzyme complex to form CO_2, acetyl-CoA, and NADH. The pyruvate dehydrogenase enzyme complex, which requires the bound cofactors thiamine pyrophosphate, lipoic acid, and FAD as well as free coenzyme A (CoASH) and NAD^+, links the TCA cycle to glycolysis.

It is known that the TCA cycle includes the following enzymes: pyruvate dehydrogenase, citrate synthase, citrate hydrolase, isocitrate dehydrogenase, oxoglutarate dehydrogenase, succinyl-CoA synthetase, succinate dehydrogenase, fumarase, malate dehydrogenase, NAD-malic enzyme and phosphoenolpyruvate carboxylase.

In particular, citrate synthase (CS) catalyses the condensation of acetyl-CoA and oxaloacetate to form the C6 molecule citrate and free CoASH, as the TCA cycle proper begins.

Malate dehydrogenase (MDH) catalyses the final step of the TCA cycle, oxidizing malate to oxaloacetate and producing NADH. This reaction catalysed by MDH is reversible, thus allowing also for the reversible reduction of oxaloacetate to malate. The enzyme MDH is important in several metabolic pathways, and higher plants contain multiple forms that differ in co-enzyme specificity and subcellular localization. Chloroplasts contain an $NADP^+$-dependent MDH that plays a critical role in balancing reducing equivalents between the cytosol and

stroma. Plants also contain NAD-dependent MDHs which are found in a) mitochondria as part of the TCA cycle; b) cytosol and peroxisomes involved in malate-aspartate shuttles; and c) glyoxisomes functioning in β-oxidation. In root nodules of nitrogen-fixing legumes, such as white clover (*Trifolium repens*) and alfalfa (*Medicago sativa*), malate serves as the primary carbon source to support the respiratory needs of the bacterial microsymbiont and the fixation of N_2 by nitrogenase, and a nodule-enhanced MDH is thus critical for nodule function.

Phosphoenolpyruvate carboxylase (PEPC) catalyses the reaction of phosphoenol-pyruvate with HCO_3^- releasing the phosphate and producing the C4 product, oxaloacetate. Oxaloacetate is commonly reduced to malate by NADH through the action of malate dehydrogenase (MDH). PEPC is a homotetrameric enzyme widely distributed in most plant tissues. In plants, PEPC fulfils various physiological roles such as the photosynthetic CO_2 fixation in C_4 and Crassulacean Acid Metabolism (CAM) plants, and the anaplerotic pathway.

While nucleic acid sequences encoding some organic acid biosynthetic enzymes have been isolated for certain species of plants, there remains a need for materials useful in modifying organic acid biosynthesis; in modifying organic acid secretion; in modifying phosphorus acquisition efficiency in plants; in modifying aluminium and acid soil tolerance in plants; in modifying nitrogen fixation and nodule function, particularly in forage legumes and grasses, including alfalfa, medics, clovers, ryegrasses and fescues, and for methods for their use.

This invention is directed towards overcoming, or at least alleviating, one or more of the difficulties or deficiencies associated with the prior art.

In one aspect, the present invention provides substantially purified or isolated nucleic acids or nucleic acid fragments encoding the organic acid biosynthetic polypeptides CS, MDH and PEPC, from a clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species, or functionally active fragments or variants of these polypeptides.

The present invention also provides substantially purified or isolated nucleic acids or nucleic acid fragments encoding amino acid sequences for a class of polypeptides which are related to CS, MDH and PEPC (from a clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species) of CS, MDH

and PEPC, or functionally active fragments or variants of CS, MDH and PEPC. Such polypeptides are referred to herein as CS-like, MDH-like and PEPC-like respectively and include polypeptides having similar functional activity.

The present invention also relates to individual or simultaneous enhancement or otherwise manipulation of CS, MDH and/or PEPC or like gene activities in plants to enhance or otherwise alter organic acid biosynthesis; to enhance or reduce or otherwise alter organic acid secretion; to enhance or reduce or otherwise alter phosphorous acquisition efficiency in plants; to enhance or reduce or otherwise alter aluminium and acid soil tolerance in plants; and/or to enhance or reduce or otherwise alter nitrogen fixation and nodule function in legumes.

The individual or simultaneous enhancement or otherwise manipulation of CS, MDH and/or PEPC or like gene activities in plants has significant consequences for a range of applications in, for example, plant production, plant performance, plant nutrition and plant tolerance. For example, it has applications in increasing plant tolerance to aluminium-toxic acid soils; in improving plant nutrient acquisition efficiency for example in increasing acquisition of phosphorus from soils; in increasing nodule function in nitrogen-fixing legumes for example leading to enhanced nitrogen fixation; in modifying the accumulation of organic acids such as citrate in fruits; in modifying the secretion of organic acids for example citrate and/or malate from plant roots.

Manipulation of CS, MDH and/or PEPC or like gene activities in plants, including legumes such as clovers (*Trifolium* species), lucerne (*Medicago sativa*) and grass species such as ryegrasses (*Lolium* species) and fescues (*Festuca* species) may be used to facilitate the production of, for example, forage legumes and forage grasses and other crops with enhanced tolerance to aluminium toxic soils; enhanced nutrient acquisition efficiency; forage legumes with enhanced nitrogen fixation; fruits with enhanced organic acid content leading to enhanced flavour and health benefits.

The clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species may be of any suitable type, including white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*), alfalfa (*Medicago sativa*), Italian or annual ryegrass (*Lolium

multiflorum), perennial ryegrass (*Lolium perenne*), tall fescue (*Festuca arundinacea*), meadow fescue (*Festuca pratensis*) and red fescue (*Festuca rubra*). Preferably the species is a clover or a ryegrass, more preferably white clover (*T. repens*) or perennial ryegrass (*L. perenne*). White clover (*Trifolium repens* L.) and perennial ryegrass (*Lolium perenne* L.) are key pasture legumes and grasses, respectively, in temperate climates throughout the world. Perennial ryegrass is also an important turf grass.

The nucleic acid or nucleic acid fragment may be of any suitable type and includes DNA (such as cDNA or genomic DNA) and RNA (such as mRNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases, and combinations thereof. The RNA is readily obtainable, for example, by transcription of a DNA sequence according to the present invention, to produce a RNA corresponding to the DNA sequence. The RNA may be synthesised *in vivo* or *in vitro* or by chemical synthesis to produce a sequence corresponding to a DNA sequence by methods well known in the art. In this specification, where the degree of sequence similarity between an RNA and DNA is such that the strand of the DNA could encode the RNA, then the RNA is said to "correspond" to that DNA.

The term "isolated" means that the material is removed from its original environment (eg. the natural environment if it is naturally occurring). For example, a naturally occurring nucleic acid or polypeptide present in a living plant is not isolated, but the same nucleic acid or polypeptide separated from some or all of the coexisting materials in the natural system, is isolated. Such an isolated nucleic acid could be part of a vector and/or such a nucleic acid could be part of a composition, and still be isolated in that such a vector or composition is not part of its natural environment. An isolated polypeptide could be part of a composition and still be isolated in that such a composition is not part of its natural environment.

By "functionally active" in respect of a nucleic acid it is meant that the fragment or variant is capable of modifying organic acid biosynthesis in a plant. A variant in this context can be an analogue, derivative or mutant and includes naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity

of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 80% identity to the functional part of the above mentioned sequence, more preferably at least approximately 90% identity, most preferably at least approximately 95% identity. Such functionally active variants and fragments include, for example, those having nucleic acid changes which result in conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least 30 nucleotides, more preferably at least 45 nucleotides, most preferably at least 60 nucleotides.

By "functionally active" in respect of a polypeptide it is meant that the fragment or variant has one or more of the biological properties of the proteins CS, CS-like, MDH, MDH-like, PEPC and PEPC-like. A variant in this context includes additions, deletions, substitutions and derivatizations of one or more of the amino acids are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 60% identity to the functional part of the above mentioned sequence, more preferably at least approximately 80% identity, most preferably at least approximately 90% identity. Such functionally active variants and fragments include, for example, those having conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least 10 amino acids, more preferably at least 15 amino acids, most preferably at least 20 amino acids.

The term "construct" as used herein refers to an artificially assembled or isolated nucleic acid molecule which includes the gene of interest. In general a construct may include the gene or genes of interest, a marker gene which in some cases can also be the gene of interest and appropriate regulatory sequences. It should be appreciated that the inclusion of regulatory sequences in a construct is optional, for example, such sequences may not be required in situations where the regulatory sequences of a host cell are to be used. The term construct includes vectors but should not be seen as being limited thereto.

The term "vector" as used herein encompasses both cloning and expression vectors. Vectors are often recombinant molecules containing nucleic acid molecules from several sources.

By "operatively linked" in respect of a regulatory element, nucleic acid or nucleic acid fragment and terminator, it meant that the regulatory element is capable of causing expression of said nucleic acid or nucleic acid fragment in a plant cell and said terminator is capable of terminating expression of said nucleic acid or nucleic acid fragment in a plant cell. Preferably, said regulatory element is upstream of said nucleic acid or nucleic acid fragment and said terminator is downstream of said nucleic acid or nucleic acid fragment.

By "an effective amount" of a nucleic acid or nucleic acid fragment it is meant an amount sufficient to result in an identifiable phenotypic trait in said plant, or a plant, plant seed or other plant part derived therefrom. Such amounts can be readily determined by an appropriately skilled person, taking into account the type of plant, the route of administration and other relevant factors. Such a person will readily be able to determine a suitable amount and method of administration. See, for example, Maniatis et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, the entire disclosure of which is incorporated herein by reference.

It will also be understood that the term "comprises" (or its grammatical variants) as used in this specification is equivalent to the term "includes" and should not be taken as excluding the presence of other elements or features.

Such nucleic acids or nucleic acid fragments could be assembled to form a consensus contig. As used herein, the term "consensus contig" refers to a nucleotide sequence that is assembled from two or more constituent nucleotide sequences that share common or overlapping regions of sequence homology. For example, the nucleotide sequence of two or more nucleic acids or nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acids or nucleic acid fragments, the sequences (and thus their corresponding nucleic acids or nucleic acid fragments) can be assembled into a single contiguous nucleotide sequence.

In a preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encodes a CS or CS-like polypeptide and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 1, 3, 4, 6, 7, 9, 99, 101, 102, 104,

114, 118 and 122 hereto (SEQ ID NOS 1, 3 to 10, 11, 13 to 16, 17, 19, 327, 329 to 335, 336, 338 to 344, 349, 351, 353 respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c); and (e) RNA sequences corresponding to the sequences recited in (a), (b), (c) and (d).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encodes a MDH or MDH-like polypeptide and including a nucleotide sequence selected from the group consisting of (a) sequence shown in Figures 11, 13, 14, 16, 17, 19, 21, 23, 25, 26, 28, 30, 31, 33, 35, 37, 38, 40, 55, 57, 58, 60, 61, 63, 64, 66, 67, 69, 70, 72, 73, 75, 76, 78, 79, 81, 82 and 84 hereto (SEQ ID NOS. 21, 23 to 29; 30, 32 to 33, 34, 36, 38, 40, 42 to 43, 44, 46, 48 to 110, 111, 113, 115, 117 to 182, 183, 185, 205, 207 to 217, 218, 220 to 251, 252, 254 to 270, 271, 273 to 275, 276, 278 to 287, 288, 290 to 292, 293, 295 to 296, 297, 299 to 301, 304 to 305, 306, 308); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c); and (e) RNA sequences corresponding to the sequences recited in (a), (b), (c) and (d).

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encodes a PEPC or PEPC-like polypeptide and including a nucleotide sequence selected from the group consisting of (a) sequences shown in Figures 42, 44, 46, 47, 49, 51, 53, 86, 88, 89, 91, 92, 94, 95, 97 and 110 hereto (SEQ ID NOS 187, 189, 191 to 197, 199, 201, 203, 310, 312 to 314, 315, 317 to 318, 319, 321 to 322, 323, 325 and 347 respectively); (b) complements of the sequences recited in (a); (c) sequences antisense to the sequences recited in (a) and (b); (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c); and (e) RNA sequences corresponding to the sequences recited in (a), (b), (c) and (d).

Nucleic acids or nucleic acid fragments encoding at least a portion of several CS, MDH and PEPC polypeptides have been isolated and identified. Genes encoding other CS or CS-like, MDH or MDH-like and PEPC or PEPC-like proteins, either as cDNAs or genomic DNAs, may be isolated directly by using all

or a portion of the nucleic acids or nucleic acid fragments of the present invention as hybridisation probes to screen libraries from the desired plant employing the methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the nucleic acid sequences of the present invention may be designed and synthesized by methods known in the art. Moreover, the entire sequences may be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labelling, nick translation, or end-labelling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers may be designed and used to amplify a part or all of the sequences of the present invention. The resulting amplification products may be labelled directly during amplification reactions or labelled after amplification reactions, and used as probes to isolate full-length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, short segments of the nucleic acids or nucleic acid fragments of the present invention may be used in protocols to amplify longer nucleic acids or nucleic acid fragments encoding homologous genes from DNA or RNA. For example, polymerase chain reaction may be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the nucleic acid sequences of the present invention, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, those skilled in the art can follow the RACE protocol (Frohman *et al.* (1988) *Proc. Natl. Acad Sci. USA* 85:8998, the entire disclosure of which is incorporated herein by reference) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Using commercially available 3' RACE and 5' RACE systems (BRL), specific 3' or 5' cDNA fragments may be isolated (Ohara *et al.* (1989) *Proc. Natl. Acad Sci USA* 86:5673; Loh *et al.* (1989) *Science* 243:217, the entire disclosures of which are incorporated herein by reference). Products generated by the 3' and 5' RACE procedures may be combined to generate full-length cDNAs.

In a further aspect of the present invention there is provided a substantially purified or isolated polypeptide from a clover (*Trifolium*), medic (*Medicago*),

ryegrass (*Lolium*) or fescue (*Festuca*) species, selected from the group consisting of CS or CS-like, MDH or MDH-like and PEPC or PEPC-like polypeptides; and functionally active fragments and variants of these polypeptides.

The clover (*Trifolium*), medic (*Medicago*), ryegrass (*Lolium*) or fescue (*Festuca*) species may be of any suitable type, including white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*), alfalfa (*Medicago sativa*), Italian or annual ryegrass (*Lolium multiflorum*), perennial ryegrass (*Lolium perenne*), tall fescue (*Festuca arundinacea*), meadow fescue (*Festuca pratensis*) and red fescue (*Festuca rubra*). Preferably the species is a clover or a ryegrass, more preferably white clover (*T. repens*) or perennial ryegrass (*L. perenne*).

In a preferred embodiment of this aspect of the invention, the substantially purified or isolated CS or CS-like polypeptide includes an amino acid sequence selected from the group consisting of sequences shown in Figures 2, 5, 8, 10, 100, 103, 115, 119, 123 hereto (SEQ ID NOS 2, 12, 18, 20, 328, 337, 350, 352 and 354 respectively); and functionally active fragments and variants thereof.

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated MDH or MDH-like polypeptide includes an amino acid sequence selected from the group consisting of sequences shown in Figures 12, 15, 18, 20, 22, 24, 27, 29, 32, 34, 36, 39, 41, 56, 59, 62, 65, 68, 71, 74, 77, 80, 83 and 85 hereto (SEQ ID NOS 22, 31, 35, 37, 39, 41, 45, 47, 112, 114, 116, 184, 186, 206, 219, 253, 272, 277, 289, 294, 297, 303, 307 and 309, respectively) and functionally active fragments and variants thereof.

In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated PEPC or PEPC-like polypeptide includes an amino acid sequence selected from the group consisting of sequences shown in Figures 43, 45, 48, 50, 52, 54, 87, 90, 93, 96, 98 and 111 hereto (SEQ ID NOS 188, 190, 198, 200, 202, 204, 311, 316, 320, 324, 326, and 348 ,respectively); and functionally active fragments and variants thereof.

In a further embodiment of this aspect of the invention, there is provided a polypeptide produced (e.g. recombinantly) from a nucleic acid or nucleic acid

fragment according to the present invention. Techniques for recombinantly producing polypeptides are known to those skilled in the art.

Availability of the nucleotide sequences of the present invention and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides may be used to immunise animals to produce polyclonal or monoclonal antibodies with specificity for peptides and/or proteins including the amino acid sequences. These antibodies may be then used to screen cDNA expression libraries to isolate full-length cDNA clones of interest.

A genotype is the genetic constitution of an individual or group. Variations in genotype are important in commercial breeding programs, in determining parentage, in diagnostics and fingerprinting, and the like. Genotypes can be readily described in terms of genetic markers. A genetic marker identifies a specific region or locus in the genome. The more genetic markers, the finer defined is the genotype. A genetic marker becomes particularly useful when it is allelic between organisms because it then may serve to unambiguously identify an individual. Furthermore, a genetic marker becomes particularly useful when it is based on nucleic acid sequence information that can unambiguously establish a genotype of an individual and when the function encoded by such nucleic acid is known and is associated with a specific trait. Such nucleic acids and/or nucleotide sequence information including single nucleotide polymorphisms (SNPs), variations in single nucleotides between allelic forms of such nucleotide sequence, may be used as perfect markers or candidate genes for the given trait.

Applicants have identified a number of SNPs of the nucleic acids or nucleic acid fragments of the present invention. These are indicated (marked with grey on the black background) in the figures that show multiple alignments of nucleotide sequences of nucleic acid fragments contributing to consensus contig sequences. See for example, Figures 3, 6, 9, 13, 16, 30, 37, 57, 60, 63, 79, 89, 92 and 104 hereto.

Accordingly, in a further aspect of the present invention, there is provided a substantially purified or isolated nucleic acid or nucleic acid fragment including a single nucleotide polymorphism (SNP) from a nucleic acid or nucleic acid fragment

according to the present invention, for example a SNP from a nucleic acid sequence shown in Figures 3, 6, 9, 13, 16, 30, 37, 57, 60, 63, 66, 67, 72, 78, 88, 94, 101 and 104 hereto; or complements or sequences antisense thereto, and functionally active fragments and variants thereof. The invention further provides a substantially purified or isolated nucleic acid or nucleic acid fragment including a single nucleotide polymorphism (SNP) isolated by the method of this invention.

In a still further aspect of the present invention there is provided a method of isolating a nucleic acid or nucleic acid fragment of the present invention including a SNP, said method including sequencing nucleic acid fragments from a nucleic acid library. The method includes the step of identifying the SNP.

The nucleic acid library may be of any suitable type and is preferably a cDNA library.

The nucleic acid or nucleic acid fragment may be isolated from a recombinant plasmid or may be amplified, for example using polymerase chain reaction.

The sequencing may be performed by techniques known to those skilled in the art.

In a still further aspect of the present invention, there is provided use of the nucleic acids or nucleic acid fragments of the present invention including SNPs, and/or nucleotide sequence information thereof, as molecular genetic markers.

In a still further aspect of the present invention there is provided use of a nucleic acid or nucleic acid fragment of the present invention, and/or nucleotide sequence information thereof, as a molecular genetic marker.

More particularly, nucleic acids or nucleic acid fragments according to the present invention and/or nucleotide sequence information thereof may be used as a molecular genetic marker for quantitative trait loci (QTL) tagging, QTL mapping, DNA fingerprinting and in marker assisted selection, particularly in clovers, alfalfa, ryegrasses and fescues. Even more particularly, nucleic acids or nucleic acid fragments according to the present invention and/or nucleotide sequence information thereof may be used as molecular genetic markers in plant improvement in relation to plant tolerance to abiotic stresses such aluminium toxic acid soils; in relation to nutrient acquisition efficiency including phosphorus; in

relation to nitrogen fixation; in relation to nodulation. Even more particularly, sequence information revealing SNPs in allelic variants of the nucleic acids or nucleic acid fragments of the present invention and/or nucleotide sequence information thereof may be used as molecular genetic markers for QTL tagging and mapping and in marker assisted selection, particularly in clovers, alfalfa, ryegrasses and fescues.

In a still further aspect of the present invention there is provided a construct or vector including a nucleic acid or nucleic acid fragment according to the present invention.

In a particularly preferred embodiment the construct or vector may include nucleic acids or nucleic acid fragments encoding both CS or CS-like and MDH or MDH-like polypeptides.

In yet another preferred embodiment the construct or vector may include nucleic acids or nucleic acid fragments encoding both MDH or MDH-like and PEPC or PEPC-like polypeptides.

In yet another preferred embodiment the construct or vector may include both CS or CS-like and PEPC or PEPC-like polypeptides.

In another preferred embodiment the construct or vector may include nucleic acids or nucleic acid fragments encoding all three of CS or CS-like, MDH or MDH-like and PEPC or PEPC-like polypeptides.

In a preferred embodiment of this aspect of the invention, the vector may include one or more regulatory element such as a promoter, one or more nucleic acids or nucleic acid fragments according to the present invention and one or more terminators; said one or more regulatory elements, one or more nucleic acids or nucleic acid fragments and one or more terminators being operatively linked.

In a preferred embodiment of the present invention the vector may contain nucleic acids or nucleic acid fragments encoding both CS or CS-like and MDH or MDH-like polypeptides, operatively linked to a regulatory element or regulatory elements, such that both CS or CS-like and MDH or MDH-like polypeptides are expressed.

In another preferred embodiment of the present invention the vector may contain nucleic acids or nucleic acid fragments encoding both CS or CS-like and PEPC or PEPC-like polypeptides, operatively linked to a regulatory element or regulatory elements, such that both CS or CS-like and PEPC or PEPC-like polypeptides are expressed.

In yet another particularly preferred embodiment of the present invention the vector or construct may contain nucleic acids or nucleic acid fragments encoding both MDH or MDH-like and PEPC or PEPC-like polypeptides, operatively linked to a regulatory element or regulatory elements, such that both MDH or MDH-like and PEPC or PEPC-like polypeptides are expressed.

In another particularly preferred embodiment of the present invention the vector may contain nucleic acids or nucleic acid fragments encoding all three of CS or CS-like, MDH or MDH-like and PEPC or PEPC-like, operatively linked to a regulatory element or regulatory elements, such that all three of CS or CS-like, MDH or MDH-like and PEPC or PEPC-like polypeptides are expressed.

The vector may be of any suitable type and may be viral or non-viral. The vector may be an expression vector. Such vectors include chromosomal, non-chromosomal and synthetic nucleic acid sequences, eg. derivatives of plant viruses; bacterial plasmids; derivatives of the Ti plasmid from *Agrobacterium tumefaciens*, derivatives of the Ri plasmid from *Agrobacterium rhizogenes*; phage DNA; yeast artificial chromosomes; bacterial artificial chromosomes; binary bacterial artificial chromosomes; vectors derived from combinations of plasmids and phage DNA. However, any other vector may be used as long as it is replicable, integrative or viable in the plant cell.

The regulatory element and terminator may be of any suitable type and may be endogenous to the target plant cell or may be exogenous, provided that they are functional in the target plant cell.

Preferably the regulatory element is a promoter. A variety of promoters which may be employed in the vectors of the present invention are well known to those skilled in the art. Factors influencing the choice of promoter include the desired tissue specificity of the vector, and whether constitutive or inducible expression is desired and the nature of the plant cell to be transformed (eg.

monocotyledon or dicotyledon). Particularly suitable constitutive promoters include the Cauliflower Mosaic Virus 35S (CaMV 35S) promoter, the maize Ubiquitin promoter, and the rice Actin promoter. Particularly suitable tissue-specific promoters include root-prevalent promoters.

A variety of terminators which may be employed in the vectors of the present invention are also well known to those skilled in the art. The terminator may be from the same gene as the promoter sequence or a different gene. Particularly suitable terminators are polyadenylation signals, such as the CaMV 35S polyA and other terminators from the nopaline synthase (*nos*) and the octopine synthase (*ocs*) genes.

The vector, in addition to the regulatory element, the nucleic acid or nucleic acid fragment of the present invention and the terminator, may include further elements necessary for expression of the nucleic acid or nucleic acid fragment, in different combinations, for example vector backbone, origin of replication (ori), multiple cloning sites, spacer sequences, enhancers, introns (such as the maize Ubiquitin *Ubi* intron), antibiotic resistance genes and other selectable marker genes [such as the neomycin phosphotransferase (*npt2*) gene, the hygromycin phosphotransferase (*hph*) gene, the phosphinothricin acetyltransferase (*bar* or *pat*) gene, the phospho-mannose isomerase (*pmi*) gene], and reporter genes (such as beta-glucuronidase (GUS) gene (*gusA*)]. The vector may also contain a ribosome binding site for translation initiation. The vector may also include appropriate sequences for amplifying expression.

As an alternative to use of a selectable marker gene to provide a phenotypic trait for selection of transformed host cells, the presence of the vector in transformed cells may be determined by other techniques well known in the art, such as PCR (polymerase chain reaction), Southern blot hybridisation analysis, histochemical GUS assays, northern and Western blot hybridisation analyses.

Those skilled in the art will appreciate that the various components of the vector are operatively linked, so as to result in expression of said nucleic acid or nucleic acid fragment. Techniques for operatively linking the components of the vector of the present invention are well known to those skilled in the art. Such techniques include the use of linkers, such as synthetic linkers, for example including one or more restriction enzyme sites.

The vectors of the present invention may be incorporated into a variety of plants, including monocotyledons (such as grasses from the genera *Lolium*, *Festuca*, *Paspalum*, *Pennisetum*, *Panicum* and other forage and turfgrasses, corn, oat, sugarcane, wheat and barley), dicotyledons (such as Arabidopsis, tobacco, clovers, medics, eucalyptus, potato, sugarbeet, canola, soybean, chickpea) and gymnosperms. In a preferred embodiment, the vectors may be used to transform monocotyledons, preferably grass species such as ryegrasses (*Lolium* species) and fescues (*Festuca* species), more preferably perennial ryegrass, including forage- and turf-type cultivars. In an alternate preferred embodiment, the vectors may be used to transform dicotyledons, preferably forage legume species such as clovers (*Trifolium* species) and medics (*Medicago* species), more preferably white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*) and alfalfa (*Medicago sativa*). Clovers, alfalfa and medics are key pasture legumes in temperate climates throughout the world.

Techniques for incorporating the vectors of the present invention into plant cells (for example by transduction, transfection or transformation) are known to those skilled in the art. Such techniques include *Agrobacterium* mediated introduction, electroporation to tissues, cells and protoplasts, protoplast fusion, injection into reproductive organs, injection into immature embryos and high velocity projectile introduction to cells, tissues, calli, immature and mature embryos. The choice of technique will depend largely on the type of plant to be transformed.

Cells incorporating the vectors of the present invention may be selected, as described above, and then cultured in an appropriate medium to regenerate transformed plants, using techniques well known in the art. The culture conditions, such as temperature, pH and the like, will be apparent to the person skilled in the art. The resulting plants may be reproduced, either sexually or asexually, using methods well known in the art, to produce successive generations of transformed plants.

In a further aspect of the present invention there is provided a plant cell, plant, plant seed or other plant part, including, e.g. transformed with, a vector, nucleic acid or nucleic acid fragment of the present invention.

The plant cell, plant, plant seed or other plant part may be from any suitable species, including monocotyledons, dicotyledons and gymnosperms. In a preferred embodiment the plant cell, plant, plant seed or other plant part may be from a monocotyledon, preferably a grass species, more preferably a ryegrass (*Lolium* species) or fescue (*Festuca* species), more preferably perennial ryegrass, including both forage- and turf-type cultivars. In an alternate preferred embodiment the plant cell, plant, plant seed or other plant part may be from a dicotyledon, preferably forage legume species such as clovers (*Trifolium* species) and medics (*Medicago* species), more preferably white clover (*Trifolium repens*), red clover (*Trifolium pratense*), subterranean clover (*Trifolium subterraneum*) and alfalfa (*Medicago sativa*).

The present invention also provides a plant, plant seed or other plant part, or a plant extract derived from a plant cell of the present invention.

The present invention also provides a plant, plant seed or other plant part, or a plant extract derived from a plant of the present invention.

In a further aspect of the present invention there is provided a method of modifying organic acid biosynthesis; of modifying organic acid secretion; of modifying phosphorous and other nutrients acquisition efficiency in plants; of modifying aluminium and acid soil tolerance in plants; of modifying nitrogen fixation and nodule function, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment according to the present invention. Preferably the nucleic acid or nucleic acid fragment is part of a vector.

Using the methods and products of the present invention, organic acid biosynthesis; organic acid secretion; phosphorous and other plant nutrient acquisition efficiency; aluminium and acid soil tolerance; nitrogen fixation and nodule function, may be increased or otherwise altered, for example by incorporating additional copies of a sense nucleic acid or nucleic acid fragment of the present invention. They may be decreased or otherwise altered, for example by incorporating an antisense nucleic acid or nucleic acid fragment of the present invention.

In a particularly preferred embodiment the method may include introducing into said plant nucleic acids or nucleic acid fragments encoding both CS or CS-like and MDH or MDH-like polypeptides.

In another preferred embodiment the method may include introducing into said plant nucleic acids or nucleic acid fragments encoding both CS or CS-like and PEPC or PEPC polypeptides.

In yet another preferred embodiment the method may include introducing into said plant nucleic acids or nucleic acid fragments encoding both MDH or MDH-like and PEPC or PEPC-like polypeptides.

In an even more preferred embodiment the method may include introducing into said plant nucleic acids or nucleic acid fragments encoding all three of CS or CS-like, MDH or MDH-like and PEPC or PEPC-like polypeptides.

The present invention will now be more fully described with reference to the accompanying Examples and drawings. It should be understood, however, that the description following is illustrative only and should not be taken in any way as a restriction on the generality of the invention described above.

In the Figures

Figure 1 shows the consensus contig nucleotide sequence of LpCSa.

Figure 2 shows the deduced amino acid sequence of LpCSa.

Figure 3 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpCSa.

Figure 4 shows the consensus contig nucleotide sequence of LpCSb.

Figure 5 shows the deduced amino acid sequence of LpCSb.

Figure 6 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpCSb.

Figure 7 shows the nucleotide sequence of LpCSc.

Figure 8 shows the deduced amino acid sequence of LpCSc.

Figure 9 shows the nucleotide sequence of LpCSd.

Figure 10 shows the deduced amino acid sequence of LpCSd.

Figure 11 shows the consensus contig nucleotide sequence of LpMDHa.

Figure 12 shows the deduced amino acid sequence of LpMDHa.

Figure 13 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpMDHa.

Figure 14 shows the consensus contig nucleotide sequence of LpMDHb.

Figure 15 shows the deduced amino acid sequence of LpMDHb.

Figure 16 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpMDHb.

Figure 17 shows the nucleotide sequence of LpMDHc.

Figure 18 shows the deduced amino acid sequence of LpMDHc.

Figure 19 shows the nucleotide sequence of LpMDHd.

Figure 20 shows the deduced amino acid sequence of LpMDHd.

Figure 21 shows the nucleotide sequence of LpMDHe.

Figure 22 shows the deduced amino acid sequence of LpMDHe.

Figure 23 shows the consensus contig nucleotide sequence of LpMDHf.

Figure 24 shows the deduced amino acid sequence of LpMDHf.

Figure 25 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpMDHf.

Figure 26 shows the nucleotide sequence of LpMDHg.

Figure 27 shows the deduced amino acid sequence of LpMDHg.

Figure 28 shows the consensus contig nucleotide sequence of LpMDHh.

Figure 29 shows the deduced amino acid sequence of LpMDHh.

Figure 30 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpMDHh.

Figure 31 shows the nucleotide sequence of LpMDHi.

Figure 32 shows the deduced amino acid sequence of LpMDHi.

Figure 33 shows the nucleotide sequence of LpMDHj.

Figure 34 shows the deduced amino acid sequence of LpMDHj.

Figure 35 shows the consensus contig nucleotide sequence of LpMDHk.

Figure 36 shows the deduced amino acid sequence of LpMDHk.

Figure 37 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpMDHk.

Figure 38 shows the nucleotide sequence of LpMDHl.

Figure 39 shows the deduced amino acid sequence of LpMDHl.

Figure 40 shows the nucleotide sequence of LpMDHm.

Figure 41 shows the deduced amino acid sequence of LpMDHm.

Figure 42 shows the nucleotide sequence of LpPEPCa.

Figure 43 shows the deduced amino acid sequence of LpPEPCa.

Figure 44 shows the consensus contig nucleotide sequence of LpPEPCb.

Figure 45 shows the deduced amino acid sequence of LpPEPCb.

Figure 46 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence LpPEPCb.

Figure 47 shows the nucleotide sequence of LpPEPCc.

Figure 48 shows the deduced amino acid sequence of LpPEPCc.

Figure 49 shows the nucleotide sequence of LpPEPCd.

Figure 50 shows the deduced amino acid sequence of LpPEPCd.

Figure 51 shows the nucleotide sequence of LpPEPCe.

Figure 52 shows the deduced amino acid sequence of LpPEPCe.

Figure 53 shows the nucleotide sequence of LpPEPCf.

Figure 54 shows the deduced amino acid sequence of LpPEPCf.

Figure 55 shows the consensus contig nucleotide sequence of TrMDHa.

Figure 56 shows the deduced amino acid sequence of TrMDHa.

Figure 57 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHa.

Figure 58 shows the consensus contig nucleotide sequence of TrMDHb.

Figure 59 shows the deduced amino acid sequence of TrMDHb.

Figure 60 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHb.

Figure 61 shows the consensus contig nucleotide sequence of TrMDHc.

Figure 62 shows the deduced amino acid sequence of TrMDHc.

Figure 63 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHc.

Figure 64 shows the consensus contig nucleotide sequence of TrMDHd.

Figure 65 shows the deduced amino acid sequence of TrMDHd.

Figure 66 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHd.

Figure 67 shows the consensus contig nucleotide sequence of TrMDHe.

Figure 68 shows the deduced amino acid sequence of TrMDHe.

Figure 69 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHe.

Figure 70 shows the consensus contig nucleotide sequence of TrMDHf.

Figure 71 shows the deduced amino acid sequence of TrMDHf.

Figure 72 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHf.

Figure 73 shows the consensus contig nucleotide sequence of TrMDHg.

Figure 74 shows the deduced amino acid sequence of TrMDHg.

Figure 75 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHg.

Figure 76 shows the consensus contig nucleotide sequence of TrMDHh.

Figure 77 shows the deduced amino acid sequence of TrMDHh.

Figure 78 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHh.

Figure 79 shows the consensus contig nucleotide sequence of TrMDHi.

Figure 80 shows the deduced amino acid sequence of TrMDHi.

Figure 81 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrMDHi.

Figure 82 shows the nucleotide sequence of TrMDHj.

Figure 83 shows the deduced amino acid sequence of TrMDHj.

Figure 84 shows the nucleotide sequence of TrMDHk.

Figure 85 shows the deduced amino acid sequence of TrMDHk.

Figure 86 shows the consensus contig nucleotide sequence of TrPEPCa.

Figure 87 shows the deduced amino acid sequence of TrPEPCa.

Figure 88 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrPEPCa.

Figure 89 shows the consensus contig nucleotide sequence of TrPEPCb.

Figure 90 shows the deduced amino acid sequence of TrPEPCb.

Figure 91 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrPEPCb.

Figure 92 shows the consensus contig nucleotide sequence of TrPEPCc.

Figure 93 shows the deduced amino acid sequence of TrPEPCc.

Figure 94 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrPEPCc.

Figure 95 shows the nucleotide sequence of TrPEPCd.

Figure 96 shows the deduced amino acid sequence of TrPEPCd.

Figure 97 shows the nucleotide sequence of TrPEPCe.

Figure 98 shows the deduced amino acid sequence of TrPEPCe.

Figure 99 shows the consensus contig nucleotide sequence of TrCSa.

Figure 100 shows the deduced amino acid sequence of TrCSa.

Figure 101 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCSa.

Figure 102 shows the consensus contig nucleotide sequence of TrCSb.

Figure 103 shows the deduced amino acid sequence of TrCSb.

Figure 104 shows the nucleotide sequences of the nucleic acid fragments contributing to the consensus contig sequence TrCSb.

Figure 105 shows the plasmid map in pGEM-T Easy of TrMDH.

Figure 106 shows the nucleotide sequence of TrMDH.

Figure 107 shows the deduced amino acid sequence of TrMDH.

Figure 108 shows the plasmid map of sense construct of TrMDH in the binary vector pPZP221:35S^2.

Figure 109 shows the plasmid map in pGEM-T Easy of TrPEPC.

Figure 110 shows the nucleotide sequence of TrPEPC.

Figure 111 shows the deduced amino acid sequence of TrPEPC.

Figure 112 shows the plasmid map of sense construct of TrPEPC in the binary vector pPZP221:35S^2.

Figure 113 shows the plasmid map in pGEM-T Easy of TrCSa.

Figure 114 shows the nucleotide sequence of TrCSa.

Figure 115 shows the deduced amino acid sequence of TrCSa.

Figure 116 shows the plasmid map of sense construct of TrCSa in the binary vector pPZP221:35S^2.

Figure 117 shows the plasmid map in pGEM-T Easy of TrCSb.

Figure 118 shows the nucleotide sequence of TrCSb.

Figure 119 shows the deduced amino acid sequence of TrCSb.

Figure 120 shows the plasmid map of sense construct of TrCSb in the binary vector pPZP221:35S^2.

Figure 121 shows the plasmid map in pGEM-T Easy of TrCSd.

Figure 122 shows the nucleotide sequence of TrCSd.

Figure 123 shows the deduced amino acid sequence of TrCSd.

Figure 124 shows the plasmid map of sense construct of TrCSd in the binary vector pPZP221:35S^2.

Figure 125 shows the plasmid maps of the modular vector system comprising a binary base vector and 7 auxiliary vectors.

Figure 126 shows an example of the modular binary transformation vector system comprising plasmid maps of the binary transformation vector backbone and 4 expression cassettes for combinatorial expression of chimeric CS and MDH and PEPC genes in auxiliary vectors (A) and the plasmid map of the T-DNA region of the final binary transformation vector (B).

Figure 127 shows the results of RT-PCR experiments performed as described in Example 6. Samples were isolated from: L, leaf; S, stolon; St, stolon tip; R, root; Rt, root tip. –C: negative (no reverse transcriptase) control; +C, positive (plasmid) control. The numbers indicate cycle numbers. A: phosphate transporter homolog; B: root iron transporter homolog.

Figure 128 shows the screening of a white clover BAC library using the phosphate transporter cDNA as a probe (A); Southern hybridisation blot of six BAC clones identified in A using the same probe (B); physical map of the phosphate transporter genomic region including the coding region and the promoter region (C).

Figure 129 shows white clover cotyledons, various stages of selection of plantlets transformed with a binary transformation vector constructed as described in Examples 4 and 5, transgenic white clover on root-inducing medium, and white clover plants transformed with genes involved in organic acid biosynthesis.

Figure 130 shows the molecular analysis of transgenic white clover plants for the presence of the chimeric MDH gene with real time PCR amplification plot and agarose gel of PCR product.

Figure 131 shows the molecular analysis of transgenic white clover plants for the presence of the chimeric PEPC gene with real time PCR amplification plot and agarose gel of PCR product.

Figure 132 shows the molecular analysis of transgenic white clover plants for the presence of the chimeric CS gene with real time PCR amplification plot and agarose gel of PCR product.

EXAMPLE 1

Preparation of cDNA libraries, isolation and sequencing of cDNAs coding for CS, CS-like, MDH, MDH-like, PEPC and PEPC-like polypeptides from white clover (*Trifolium repens*) and perennial ryegrass (*Lolium perenne*)

cDNA libraries representing mRNAs from various organs and tissues of white clover (*Trifolium repens*) and perennial ryegrass (*Lolium perenne*) were prepared. The characteristics of the white clover and perennial ryegrass libraries, respectively, are described below (Tables 1 and 2).

TABLE 1

cDNA libraries from white clover (*Trifolium repens*)

Library	Organ/Tissue
01wc	Whole seedling, light grown
02wc	Nodulated root 3, 5, 10, 14, 21 &28 day old seedling
03wc	Nodules pinched off roots of 42 day old rhizobium inoculated white clover
04wc	Nodulated white clover cut leaf and stem collected after 0, 1, 4, 6 &14 h after cutting
05wc	Non-nodulated Inflorescences: <50% open, not fully open and fully open
06wc	Dark grown etiolated
07wc	Inflorescence – very early stages, stem elongation, < 15 petals, 15-20 petals
08wc	seed frozen at –80°C, imbibed in dark overnight at 10°C
09wc	Drought stressed plants
10wc	AMV infected leaf
11wc	WCMV infected leaf

Library	Organ/Tissue
12wc	Phosphorus starved plants
13wc	Vegetative stolon tip
14wc	stolon root initials
15wc	Senescing stolon
16wc	Senescing leaf

TABLE 2

cDNA libraries from perennial ryegrass (*Lolium perenne*)

Library	Organ/Tissue
01rg	Roots from 3-4 day old light-grown seedlings
02rg	Leaves from 3-4 day old light-grown seedlings
03rg	Etiolated 3-4 day old dark-grown seedlings
04rg	Whole etiolated seedlings (1-5 day old and 17 days old)
05rg	Senescing leaves from mature plants
06rg	Whole etiolated seedlings (1-5 day old and 17 days old)
07rg	Roots from mature plants grown in hydroponic culture
08rg	Senescent leaf tissue
09rg	Whole tillers and sliced leaves (0, 1, 3, 6, 12 and 24 h after harvesting)
10rg	Embryogenic suspension-cultured cells
11rg	Non-embryogenic suspension-cultured cells
12rg	Whole tillers and sliced leaves (0, 1, 3, 6, 12 and 24 h after harvesting)
13rg	Shoot apices including vegetative apical meristems
14rg	Immature inflorescences including different stages of inflorescence

Library	Organ/Tissue
	meristem and inflorescence development
15rg	Defatted pollen
16rg	Leaf blades and leaf sheaths (*rbcL*, *rbcS*, *cab*, *wir2A* subtracted)
17rg	Senescing leaves and tillers
18rg	Drought-stressed tillers (pseudostems from plants subjected to PEG-simulated drought stress)
19rg	Non-embryogenic suspension-cultured cells subjected to osmotic stress (grown in media with half-strength salts) (1, 2, 3, 4, 5, 6, 24 and 48 h after transfer)
20rg	Non-embryogenic suspension-cultured cells subjected to osmotic stress (grown in media with double-strength salts) (1, 2, 3, 4, 5, 6, 24 and 48 h after transfer)
21rg	Drought-stressed tillers (pseudostems from plants subjected to PEG-simulated drought stress)
22rg	Spikelets with open and maturing florets
23rg	Mature roots (specific subtraction with leaf tissue)

The cDNA libraries may be prepared by any of many methods available. For example, total RNA may be isolated using the Trizol method (Gibco-BRL, USA) or the RNeasy Plant Mini kit (Qiagen, Germany), following the manufacturers' instructions. cDNAs may be generated using the SMART PCR cDNA synthesis kit (Clontech, USA), cDNAs may be amplified by long distance polymerase chain reaction using the Advantage 2 PCR Enzyme system (Clontech, USA), cDNAs may be cleaned using the GeneClean spin column (Bio 101, USA), tailed and size fractionated, according to the protocol provided by Clontech. The cDNAs may be introduced into the pGEM-T Easy Vector system 1 (Promega, USA) according to the protocol provided by Promega. The cDNAs in the pGEM-T Easy plasmid vector are transfected into *Escherichia coli* Epicurian coli XL10-Gold

ultra competent cells (Stratagene, USA) according to the protocol provided by Stratagene.

Alternatively, the cDNAs may be introduced into plasmid vectors for first preparing the cDNA libraries in Uni-ZAP XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA, USA). The Uni-ZAP XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut pBluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into *E. coli* DH10B cells according to the manufacturer's protocol (GIBCO BRL Products).

Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Plasmid DNA preparation may be performed robotically using the Qiagen QiaPrep Turbo kit (Qiagen, Germany) according to the protocol provided by Qiagen. Amplified insert DNAs are sequenced in dye-terminator sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"). The resulting ESTs are analysed using an Applied Biosystems ABI 3700 sequence analyser.

EXAMPLE 2

DNA sequence analyses

The cDNA clones encoding CS, CS-like, MDH, MDH-like, PEPC and PEPC-like polypeptides were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul *et al.* (1993) *J. Mol. Biol.* 215:403-410) searches. The cDNA sequences obtained were analysed for similarity to all publicly available DNA sequences contained in the *e*Bioinformatics nucleotide database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the SWISS-PROT protein sequence database using BLASTx algorithm (v 2.0.1) (Gish and States (1993) *Nature Genetics* 3:266-272) provided by the NCBI.

The cDNA sequences obtained and identified were then used to identify additional identical and/or overlapping cDNA sequences generated using the BLASTN algorithm. The identical and/or overlapping sequences were subjected to a multiple alignment using the CLUSTALw algorithm, and to generate a consensus contig sequence derived from this multiple sequence alignment. The consensus contig sequence was then used as a query for a search against the SWISS-PROT protein sequence database using the BLASTx algorithm to confirm the initial identification.

EXAMPLE 3

Identification and full-length sequencing of cDNAs encoding CS, MDH and PEPC polypeptides

To fully characterise for the purposes of the generation of probes for hybridisation experiments and the generation of transformation vectors, a set of cDNAs encoding white clover CS, MDH and PEPC polypeptides was identified and fully sequenced.

Full-length cDNAs were identified from our EST sequence database using relevant published sequences (NCBI databank) as queries for BLAST searches. Full-length cDNAs were identified by alignment of the query and hit sequences using Sequencher (Gene Codes Corp., Ann Arbor, MI 48108, USA). The original plasmid was then used to transform chemically competent XL-1 cells (prepared in-house, $CaCl_2$ protocol). After colony PCR (using HotStarTaq, Qiagen) a minimum of three PCR-positive colonies per transformation were picked for initial sequencing with M13F and M13R primers. The resulting sequences were aligned with the original EST sequence using Sequencher to confirm identity and one of the three clones was picked for full-length sequencing, usually the one with the best initial sequencing result.

Sequencing of all cDNAs was completed by primer walking, i.e. oligonucleotide primers were designed to the initial sequence obtained using M13F and M13R oligonucleotide primers and used for further sequencing. The sequences of the oligonucleotide primers are shown in Table 2.

Contigs were then assembled in Sequencher. The contigs include the sequences of the SMART primers used to generate the initial cDNA library as well

as pGEM-T Easy vector sequence up to the EcoRI cut site both at the 5' and 3' end.

Plasmid maps and the full cDNA sequences of TrCSa, TrCSb, TrCSd, TrMDH and TrPEPC polypeptides were obtained (Figures 1, 2, 5, 6, 9, 10, 13, 14, 17, 18, 21, 22, 25, 26, 29 and 30).

TABLE 3

List of primers used for sequencing of the full-length cDNAs encoding CS, MDH and PEPC

gene name	clone ID	sequencing primer	primer sequence (5'>3')
TrCSa	05wc1HsB08	05wc1HsB08.f1	TTGCCCGAGGCTATACTGTGGC
		05wc1HsB08.f2	CAGCTCACCTAGTTGCTAG
		05wc1HsB08.f3	CCATGGCCTAATGTTGATGC
		05wc1HsB08.r1	TTGGCCTTTCAAGTGGCATTCC
		05wc1HsB08.r2	CAGAATGGGAGGCACGACTTC
		05wc1HsB08.r3	ATGTGAGCATAGTTTGCACC
TrCSb	05wc2HsD09	05wc2HsD09.f1	GACTGCCAGAAAACACTTCCAGG
		05wc2HsD09.f2	ATGACTGCTTTAGTGTGG
		05wc2HsD09.r1	CTCAAGTTTCTCCAGTGTGACAC
		05wc2HsD09.r2	TGACTTATGTATCCCACC
		05wc2HsD09.r3	GCTCTGAATGGTTTAGCTGG
TrCSd	10wc1BsF10	10wc1BsF10.f1	GCACTGCCTGTTTCTGCTCATCC
		10wc1BsF10.f2	AGCCAACTTATGAGGATAGC
		10wc1BsF10.r1	CTCCAATACTCCTCGCGACGCC
		10wc1BsF10.r2	AGGCACAACCTGGCCACTG
		10wc1BsF10.r3	ACGTTGCCACCTTCATGATC
TrMDH	13wc1NsD01	13wc1NsD01.f1	GTTGTTATACCTGCTGGTGTT
		13wc1NsD01.r1	CTCACTCAACCCTTGGAGAT
TrPEPC	15wc1DsH12	15wc1DsH12.f1	TCCTAAGAAACTTGAAGAGCTCGG
		15wc1DsH12.f2	AGATGTTTGCTTACTAGC
		15wc1DsH12.r1	GCCAGCAGCAATACCCTTCATGG
		15wc1DsH12.r2	TTGCTTCTCAACTGTTCC

EXAMPLE 4

Development of binary transformation vectors containing chimeric genes with cDNA sequences encoding CS, MDH and PEPC

To alter the expression of the polypeptides involved in organic acid biosynthesis to improve phosphorus acquisition efficiency as well as aluminium and acid soil tolerance in forage plants, a set of sense binary transformation vectors was produced.

The pPZP221 binary transformation vector (Hajdukiewicz *et al.*, 1994) was modified to contain the $35S^2$ cassette from pKYLX71:$35S^2$ (Schardl *et al.*, 1987) as follows: pKYLX71:$35S^2$ was cut with ClaI. The 5' overhang was filled in using Klenow and the blunt end was A-tailed with Taq polymerase. After cutting with EcoRI, the 2kb fragment with an EcoRI-compatible and a 3'-A tail was gel-purified. pPZP221 was cut with HindIII and the resulting 5' overhang filled in and T-tailed with Taq polymerase. The remainder of the original pPZP221 multi-cloning site was removed by digestion with EcoRI, and the expression cassette cloned into the EcoRI site and the 3' T overhang restoring the HindIII site. This binary vector contains between the left and right border the plant selectable marker gene aacC1 under the control of the 35S promoter and 35S terminator and the pKYLX71:$35S^2$-derived expression cassette with a CaMV 35S promoter with a duplicated enhancer region and an rbcS terminator.

A GATEWAY® cloning cassette (Invitrogen) was introduced into the multicloning site of the pPZP221:$35S^2$ vector obtained as described following the manufacturer's protocol.

cDNA fragments were generated by high fidelity PCR with a proofreading DNA polymerase using the original pGEM-T Easy plasmid cDNA as a template. The primers used (Table 3) contained *att*B sequences for use with recombinases utilising the GATEWAY® system (Invitrogen). The resulting PCR fragments were used in a recombination reaction with pDONR® vector (Invitrogen) to generate entry vectors. In a further recombination reaction, the cDNAs encoding the open reading frame sequences were transferred from the entry vector to the GATEWAY®-enabled pPZP221:$35S^2$ vector.

The orientation of the constructs (sense or antisense) was checked by restriction enzyme digest and sequencing which also confirmed the correctness of the sequence. Transformation vectors containing chimeric genes using full-length open reading frame cDNAs encoding white clover TrCSa, TrCSb, TrCSd, TrMDH and TrPEPC proteins in sense orientation under the control of the CaMV $35S^2$ promoter were generated (Figures 4, 8, 12, 16, 20, 24, 28 and 32).

TABLE 4

List of primers used to PCR-amplify the open reading frames of cDNAs encoding CS, MDH and PEPC

gene name	clone ID	primer	primer sequence (5'>3')
TrCSa	05wc1HsB08	05wc1HsB08f	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGATCTTAATGGCGTTCTTTCG
		05wc1HsB08r	GGGGACCACTTTGTACAAGAAAGCTGGGTTTTCAATTTTAGGACGATGCG
TrCSb	05wc2HsD09	05wc2HsD09f	GGGGACAAGTTTGTACAAAAAAGCAGGCTTTGTTGATTGATCTTAATGGC
		05wc2HsD09r	GGGGACCACTTTGTACAAGAAAGCTGGGTTAGTAATCCACAGATAACCG
TrCSd	10wc1BsF10	10wc1BsF10f	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTAGATTGTTGATTGATCTAAATGGC
		10wc1BsF10r	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGATTCAATTTTAGGATGATGCACC
TrMDH	13wc1NsD01	13wc1NsD01f	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTAGAAATTCCCATTACCATTCATTCC
		13wc1NsD01r	GGGGACCACTTTGTACAAGAAAGCTGGGTCTAGATTGACATTCTCTCGCATGGACGC
TrPEPC	15wc1DsH12	15wc1DsH12f	GGGGACAAGTTTGTACAAAAAAGCAGGCTTGAGAAGGAGTGAATTGCTCC
		15wc1DsH12r	GGGGACCACTTTGTACAAGAAAGCTGGGTATGATATCTTAGCACACACTTAAC

EXAMPLE 5

Development of binary transformation vectors containing chimeric genes with a combination of 2 or more cDNA sequences encoding CS, MDH and PEPC

To alter the expression of the polypeptides involved in organic acid biosynthesis to improve phosphorus acquisition efficiency as well as aluminium

and acid soil tolerance in forage plants, a modular binary transformation vector system was used (Figure 125). The modular binary vector system enables simultaneous integration of up to seven transgenes the expression of which is controlled by individual promoter and terminator sequences into the plant genome (Goderis *et al.*, 2002).

The modular binary vector system consists of a pPZP200-derived vector (Hajdukiewicz *et al.*, 1994) backbone containing within the T-DNA a number of unique restriction sites recognised by homing endonucleases. The same restriction sites are present in pUC18-based auxiliary vectors flanking standard multicloning sites. Expression cassettes comprising a selectable marker gene sequence or a cDNA sequence to be introduced into the plant under the control of regulatory sequences like promoter and terminator can be constructed in the auxiliary vectors and then transferred to the binary vector backbone utilising the homing endonuclease restriction sites. Up to seven expression cassettes can thus be integrated into a single binary transformation vector. The system is highly flexible and allows for any combination of cDNA sequence to be introduced into the plant with any regulatory sequence.

For example, a selectable marker cassette comprising the nos promoter and nos terminator regulatory sequences controlling the expression of the nptII gene was PCR-amplified using a proofreading DNA polymerase from the binary vector pKYLX71:35S[2] and directionally cloned into the AgeI and NotI sites of the auxiliary vector pAUX3166. Equally, other selectable marker cassettes can be introduced into any of the auxiliary vectors.

In another example, the expression cassette from the binary vector pWM5 consisting of the ASSU promoter and the tob terminator was PCR-amplified with a proofreading DNA polymerase and directionally cloned into the AgeI and NotI sites of the auxiliary vector pAUX3169. Equally, other expression cassettes can be introduced into any of the auxiliary vectors.

In yet another example, the expression cassette from the direct gene transfer vector pDH51 was cut using EcoRI and cloned directly into the EcoRI site of the auxiliary vector pAUX3132.

TABLE 5

List of primers used to PCR-amplify plant selectable marker cassettes and the regulatory elements used to control the expression of CS, MDH and PEPC genes

expression cassette	primer	primer sequence (5'>3')
nos::nptII-nos	forward	ATAATAACCGGTTGATCATGAGCGGAGAATTAAGGG
	reverse	ATAATAGCGGCCGCTAGTAACATAGATGACACCGCG
35S::aacC1-35S	forward	AATAGCGGCCGCGATTTAGTACTGGATTTTGG
	reverse	AATAACCGGTACCCACGAAGGAGCATCGTGG
$35S^2$::rbcS	forward	ATAATAACCGGTGCCCGGGGATCTCCTTTGCC
	reverse	ATAATAGCGGCCGCATGCATGTTGTCAATCAATTGG
assu::tob	forward	TAATACCGGTAAATTTATTATGRGTTTTTTTCCG
	reverse	TAATGCGGCCGCTAAGGGCAGCCCATACAAATGAAGC

The expression cassettes were further modified by introducing a GATEWAY® cloning cassette (Invitrogen) into the multicloning site of the respective pAUX vector following the manufacturer's protocol. In a recombination reaction, the cDNAs encoding the open reading frame sequences were transferred from the entry vector obtained as described in Example 4 to the GATEWAY®-enabled pAUX vector. Any combination of the regulatory elements with cDNA sequences of TrCSa, TrCSb, TrCSd, TrMDH and TrPEPC can be obtained. One typical example is given in Figure 126 with expression cassettes comprising the nptII plant selectable marker, TrPEPC, TrCSa and TrMDH.

Complete expression cassettes comprising any combination of regulatory elements and cDNA sequences to be introduced into the plant were then cut from the auxiliary vectors using the respective homing endonuclease and cloned into the respective restriction site on the binary vector backbone. After verification of the construct by nucleotide sequencing, the binary transformation vector

comprising a number of expression cassettes was used to generate transgenic white clover plants.

EXAMPLE 6

Isolation of regulatory elements to direct expression of chimeric genes encoding CS, MDH and PEPC involved in organic acid biosynthesis

To direct the expression of chimeric white clover genes TrCSa, TrCSb, TrCSd, TrMDH and TrPEPC involved in organic acid biosynthesis to specific tissues, regulatory elements showing specificity for expression in root or root tip tissue were identified and isolated.

Using the BLASTn algorithm, white clover EST sequence collections prepared as detailed in Examples 1 and 2 were searched with nucleotide sequences representing genes with known root-specific expression identified in GenBank as queries. Suitable candidate ESTs were identified and oligonucleotide primers for reverse transcription-PCR (RT-PCR) were designed (see Table 4).

TABLE 6

Oligonucleotide primers used in reverse transcription-PCR to confirm tissue specificity of candidate white clover ESTs

gene	forward primer (5'->3')	reverse primer (5'->3')
histone (internal control)	CCGATTCCGTTTCAATGGCTCGTA	GCCATCCTTAACCCTAAGCACGT
white clover phosphate transporter homolog	TTGCATTTGCTTGGAACAACTAG	GCAAGAGCAAACATGAAACCA
white clover root iron transporter homolog	ATGGGTCTTGGTGGTTGCA	GCAGCAAGAAGATCAACCAAAGCCA

Total RNA for RT-PCR experiments was isolated from a leaf, stolon, stolon tip, root and root tip of white clover plants grown in the glasshouse using the TRIZOL method. Reverse transcription was performed using SuperScriptII (Invitrogen) following the supplier's instructions. Preliminary PCR reactions using Dynazyme as the DNA polymerase were set up to determine the correct amount of template using the PCR primers for the internal control (histone). The results of this preliminary PCR were used to set up another round of PCR to determine the optimum number of cycles for linear amplification. The final PCR amplifications

were performed using the following cycling conditions: 94 °C, 4 min., 1 time; 94 °C, 15 sec., 60 °C, 30 sec., 72 °C, 2 min., x times; 72 °C, 10 min., 1 time. The number of cycles in the amplification (x) was found to be dependent on the relative abundance of transcript and had to be optimised for each template.

RT-PCR results using a white clover histone gene as an internal constitutively expressed control confirmed the tissue-specificity of two candidate ESTs to be root-prevalent (Figure 127 A and B). These were a phosphate transporter homolog (clone name 02wc1DsG07) and a root iron transporter homolog (clone name 05wc1IsB11).

A spotted white clover BAC library consisting of 50,304 clones with an estimated 99% genome coverage (6.3 genome equivalents) was screened using the phosphate transporter homolog EST nucleotide sequence as a probe. A number of positive BAC clones could be identified (Figure 128 A). After Southern hybridisation blotting (Figure 128 B) a 7.5 kb EcoRV genomic DNA fragment was selected and fully sequenced. The fragment contained the phosphate transporter homolog open reading frame and 4 kb of upstream sequence including the promoter region. A physical map of the genomic DNA fragment including the promoter region is shown in Figure 128 C.

EXAMPLE 7

Production by *Agrobacterium*-mediated transformation and analysis of transgenic white clover plants carrying chimeric genes encoding CS, MDH and PEPC involved in organic acid biosynthesis

A set of binary transformation vectors carrying chimeric white clover genes to alter the expression of the polypeptides involved in organic acid biosynthesis to improve phosphorus acquisition efficiency as well as aluminium and acid soil tolerance in forage plants were produced as detailed in Examples 4 and 5.

Agrobacterium-mediated gene transfer experiments were performed using these transformation vectors.

The production of transgenic white clover plants carrying the white clover TrCSa, TrCSb, TrCSd, TrMDH and TrPEPC cDNAs, either singly or in combination, is described here in detail (Table 7).

Preparation of *Agrobacterium*

Agrobacterium tumefaciens strain AGL-1 transformed with one of the binary vector constructs detailed in Example 6 were streaked on LB medium containing 50 µg/ml rifampicin and 50 µg/ml kanamycin and grown at 27 °C for 48 hours. A single colony was used to inoculate 5 ml of LB medium containing 50 µg/ml rifampicin and 50 µg/ml kanamycin and grown over night at 27 °C and 250 rpm on an orbital shaker. The overnight culture was used as an inoculum for 500 ml of LB medium containing 50 µg/ml kanamycin only. Incubation was over night at 27 °C and 250 rpm on an orbital shaker in a 2 l Erlenmeyer flask.

Preparation of white clover seeds

1 spoon of seeds (ca. 500) was placed into a 280 µm mesh size sieve and washed for 5 min under running tap water, taking care not to wash seeds out of sieve. In a laminar flow hood, seeds were transferred with the spoon into an autoclaved 100 ml plastic culture vessel. A magnetic stirrer (wiped with 70% EtOH) and ca. 30 ml 70% EtOH were added, and the seeds were stirred for 5 min. The EtOH was discarded and replaced by 50 ml 1.5% sodium hypochlorite. The seeds were stirred for an additional 45 - 60 min on a magnetic stirrer. The sodium hypochlorite was then discarded and the seeds rinsed 3 to 4 times with autoclaved ddH_2O. Finally 30 ml of ddH_2O were added, and seeds incubated over night at 10 - 15°C in an incubator.

***Agrobacterium*-mediated transformation of white clover**

The seed coat and endosperm layer of the white clover seeds prepared as above were removed with a pair of 18 G or 21 G needles. The cotyledons were cut from the hypocotyl leaving a ca. 1.5 mm piece of the cotyledon stalk. The cotyledons were transferred to a petridish containing ddH_2O. After finishing the isolation of clover cotyledons, ddH_2O in the petridish was replaced with *Agrobacterium* suspension (diluted to an OD_{600} = 0.2 - 0.4). The petridish was sealed with its lid and incubated for 40 min at room temperature.

After the incubation period, each cotyledon was transferred to paper towel using the small dissection needle, dried and plated onto regeneration medium RM73. The plates were incubated at 25°C with a 16h light/8h dark photoperiod. On day 4, the explants were transferred to fresh regeneration medium. Cotyledons

transformed with *Agrobacterium* were transferred to RM73 containing cefotaxime (antibacterial agent) and gentamycin. The dishes were sealed with Parafilm and incubated at 25°C under a 16/8 h photoperiod. Explants were subcultured every three weeks for a total of nine weeks onto fresh RM 73 containing cefotaxime and gentamycin. Shoots with a green base were then transferred to root-inducing medium RIM. Roots developed after 1 – 3 weeks, and plantlets were transferred to soil when the roots were well established.

Preparation of genomic DNA for real-time PCR and analysis for the presence of transgenes

3 – 4 leaves of white clover plants regenerated on selective medium were harvested and freeze-dried. The tissue was homogenised on a Retsch MM300 mixer mill, then centrifuged for 10 min at 1700xg to collect cell debris. Genomic DNA was isolated from the supernatant using Wizard Magnetic 96 DNA Plant System kits (Promega) on a Biomek FX (Beckman Coulter). 5 µl of the sample (50 µl) were then analysed on an agarose gel to check the yield and the quality of the genomic DNA.

Genomic DNA was analysed for the presence of the transgene by real-time PCR using SYBR Green chemistry. PCR primer pairs were designed using MacVector (Accelrys) or PrimerExpress (ABI). The forward primer was located within the $35S^2$ promoter region and the reverse primer within the transgene to amplify products of approximately 150 - 250 bp as recommended. The positioning of the forward primer within the $35S^2$ promoter region guaranteed that endogenous genes in white clover were not detected.

5 µl of each genomic DNA sample was run in a 50 µl PCR reaction including SYBR Green on an ABI 7700 (Applied Biosystems) together with samples containing DNA isolated from wild type white clover plants (negative control), samples containing buffer instead of DNA (buffer control) and samples containing the plasmid used for transformation (positive plasmid control). Cycling conditions used were 2 min. at 50 °C, 10 min. at 95 °C and then 40 cycles of 15 sec. at 95 °C, 1 min. at 60 °C.

Preparation of genomic DNA and analysis of DNA for presence and copy number of transgene by Southern hybridisation blotting

Genomic DNA for Southern hybridisation blotting was obtained from leaf material of white clover plants following the CTAB method. Southern hybridisation blotting experiments were performed following standard protocols as described in Sambrook *et al.* (1989). In brief, genomic DNA samples were digested with appropriate restriction enzymes and the resulting fragments separated on an agarose gel. After transfer to a membrane, a cDNA fragment representing a transgene or selectable marker gene was used to probe the size-fractionated DNA fragments. Hybridisation was performed with either radioactively labelled probes or using the non-radioactive DIG labelling and hybridisation protocol (Boehringer) following the manufacturer's instructions.

Plants were obtained after transformation with all chimeric constructs and selection on medium containing gentamycin. Details of plant analysis are given in Table 5 and Figures 130, 131 and 132.

TABLE 7

Transformation of white clover with binary transformation vectors comprising cDNAs encoding CS, MDH and PEPC involved in organic acid biosyntheses, selection and molecular analysis of regenerated plants.

construct	cotyledons transformed	selection into RIM	soil	QPCR-positive	Southern	copy number range
pPZP221-35S2::TrMDH	2739	72	45	43	n/d	
pPZP221-35S2::TrCS	2550	39	7	nd	n/d	
pPZP221-35S2::TrPEPC	2730	44	10	nd	n/d	

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. **215**, 403-410.

Frohman, M.A., Dush, M.K., Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998.

Gish, W., States, D.J. (1993) Identification of protein coding regions by database similarity search. Nature Genetics **3**, 266-272.

Goderis, I., De Bolle, M.F.C., Francois, I., Wouters, P.F.J., Broekaert, W.F., and Cammue, B.P.A. (2002) A set of modular plant transformation vectors allowing flexible insertion of up to six expression units. Plant Molecular Biology **50**, 17-27.

Hajdukiewicz P, Svab Z, Maliga P. (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Mol Biol. **25**, 989-94.

Loh, E.Y., Elliott, J.F., Cwirla, S., Lanier, L.L., Davis, M.M. (1989). Polymerase chain reaction with single-sided specificity: Analysis of T-cell receptor delta chain. *Science* **243**, 217-220.

Ohara, O., Dorit, R.L., Gilbert, W. (1989). One-sided polymerase chain reaction: The amplification of cDNA. *Proc. Natl. Acad Sci USA* **86**, 5673-5677

Sambrook, J., Fritsch, E.F., Maniatis, T. (1989). Molecular Cloning. A Laboratory Manual. Cold Spring Harbour Laboratory Press

Schardl, C.L., Byrd, A.D., Benzion, G., Altschuler, M.A., Hildebrand, D.F., Hunt, A.G. (1987) Design and construction of a versatile system for the expression of foreign genes in plants. Gene **61**, 1-11

Finally, it is to be understood that various alterations, modifications and/or additions may be made without departing from the spirit of the present invention as outlined herein.